

CrossOSN-U: Hetero

This sub-dataset consists of two sql tables for Twitter users and two sql tables for YouTube users. The overlapped users are identified from the ‘**globalUserId**’ attribute.

➤ Twitter tables:

Table 1 – twitter user info table (*twitteruserinfo_overlapped.sql*, *twitteruserinfo_nonoverlapped.sql*)

This table contains the basic user profiles for all the downloaded Twitter users, as well as their social relationships (i.e. following and be followed). One is about the 11,687 overlapped users with information on both OSNs, and the other is the 27,972 non-overlapped users with information only on Twitter. A snapshot of this table is provided as below:

globalUserId	twitterUserId	location	description	friendsCount	followersCount	favoriteCount	statusCount	friendIdList	followerIdList
100000264454720995534	384075177	Detroit, MI	Family, Work, Future, Slim Sha	109	37	3	478	16409683,1242979369,18793	2229282720,1242979369,187
100000156269872029369	242698957	Boulder, Colorado, USA	I'm a human-powered urbanist	38	85	0	381	1313733494,132761592,5095	28845755,1378220179,11131
1000007524255117158360	10054462	Portland, OR		83	47	1	323	66316550,1113103951,10937	1418236736,22220563,17884
1000007788736839028580	14202357	upstate Ohio	Counterpoints the surrealism o	290	515	752	31363	106170894,10483527,426540	1653067303,546102159,304
100000876457024883045	16263710	Cary and Raleigh, NC	software developer, father, bx	384	169	556	1329	12524622,45431392,569858	235414297,1267811592,8114
10001231718417075017	14639823			151	60	17	864	1088175470,1638960320,800	45194439,1706741066,11175
100013687611442258419	75245022	Exeter, Devon	We're travel insurance speciali	514	1139	21	5730	901662884,308843669,88995	150433,901662884,19137057
100016383867666174158	15659711	33720, Finland / PE28, UK	Software developer, webmasti	295	240	163	5635	1660468830,62513246	39103,16326088,219612524,22040
100018122474953895346	45590048	Toronto	Communications architect and	111	195	0	246	118082023,155819788,11302	37704492,2233709426,35553
100018280541721034027	208746832	Zwolle	zvollenear 17 jaar start-uptor	215	160	7	238	563096286,2168320219,1312	1172666534,60921647,1138
100019916313199822340	196278421	Based in Fresno, Ca	Helping Independent, Entertai	2012	1207	3	1652	338650840,11609972,599240	257883956,63328696,29975
100021071940321200632	282116170	Phoenix, AZ	Human, My Daughters Father,	1996	1564	16135	6451	37740430,15258479,508310	308997977,2188515541,152
10002273349715143735	15566816	Maryland	#Business, #InternetMarketing	284	300	114	2791	1392163634,149307285,1271	52083746,56488709,3126
100023202995607020450	147938251	Carmel, IN	Commercial and Residential hui	15	15	0	63	172029210,436223979,31213	2219486950,1022736516,126
100024084790537014786	69945763	Fort Worth, Texas	on my blog	509	359	20	3303	20845950,57919344,982800	59063097,20845950,7693481
100024316682214673602	50149573	Waterloo, Ontario	Hot sauce enthusiast, geek, qu	1284	250	4	1573	19989713,16076032,1031441	253263664,17809248,332462
10003047359163401277	163436056		Il mio giorno è domani. Oggi st	517	167	221	2795	306882116,1707667939,2026	1707667939,306882116,330
100030967570282228843	14529220	27.971988,-82.560018	Social media strategist, intrap	1467	1487	10	10203	87454923020,1904768552,	43124468,216370193,43311
10003388837734212705	20897400		Passionné par les NTIC / Réalit	1080	1250	7	10969	914409992,377915374,76175	377915374,914409992,76175
100034066616152071873	10623142	Denver, Colorado	Author of <i>The Mobile Marketing</i>	3487	7133	380	7202	855128708,139823781,28997	431109210,196330518,7380
10003940815814271976311	52873477	Copenhagen, Berlin	Danish electronic singer-song	334	154	99	970	1334819904,1842181914,169	2219280278,22328017,1842
100039580744219512618	1071701		Developer of Smart Grid deplo	351	430	225	7917	14749606,52374731,148633	1871905105,2183079618,161
10003942828553575272	14612298	Buckinghamshire, UK	Blogger, writer, speaker & love	26765	40832	1244	24039	250006522,16399885,27700	15202199,14099,47083
1000395792233109552669	31665547	Reno, NV	I'm DeAño, aka EvaAnubis. I'm	354	261	99	3823	161181903,3979068,282194	1644718472,28219407,57663
10003964993214603355	14331818	Silver Spring, MD	Husband, Dad, Web Developer	484	2570	2264	14398	573206556,16453846,1851086	14565042,1916754971,16022
10003790469839857326	325286981	Midwest	RH-Entrepreneur-Writer-Smart	360	257	3	1270	170342114,17790274,19212	20149247,15912508,1703042
100038801356570551641	697843	Mountain View, CA	Senior Software Engineer for h	647	834	59	7908	96206175,3419450,1426770	105591152,1897358886,6362
100040169624923699156	17623660	Holtsville, NY		63	22	2	150	319561543,35406821,355906	15374268,43201056,3666108
10004089516481379307	7532972	Miamisburg, OH	Poly & opinionated, musician, e	95	203	134	1413	89496688,1434735,1379657	60945281,414779076,330807

Figure 1: A snapshot of twitter user info table

Each row represents the user profiles of one user, ‘globalUserId’ is used to identify the overlapped users in both OSNs, ‘friendIdList’ and ‘followerIdList’ contain all the Twitter user ids of this user’s friends and followers, respectively. Note that the Twitter user ids are separated by the ‘,’ separator.

Table 2 – twitter users’ LDA-based tweeting features

(*twitterusertweetLDAfeatures_overlapped.theta*, *twitterusertweetLDAfeatures_nonoverlapped.theta*)

This table consists of the users’ tweeting features via the standard LDA model (txt format), where the topic number is 60. Also, one is for the overlapped users and the other is for the non-overlapped users. Each row in this table represents the tweeting features of one user, the feature number is 60. The users order is the same as that in table 1.

➤ YouTube tables:

Table 3 – YouTube user info table (*youtubeuserinfo_overlapped.sql*, *youtubeuserinfo_nonoverlapped.sql*)

This table contains the basic user profiles for all the downloaded YouTube users, as well as their video-related behaviors (i.e. uploaded, favorite and add-to-playlist). Similar to table 1 of Twitter tables, one is about the overlapped users and the other is for the non-overlapped users. A snapshot of this table is presented as below:

globalUserId	location	uploadedVideos	favoriteVideos	playListVideos	videoCount
100000264454720995534	US	wRMLGxGhzJU,Wf28OCzr1qk,	ooOL4T-BAg0,LOJu03a18eI,fy	eTr7UtnUfM,	19
100001562698720292969	US	jtCM5cj3mlc,	Xj8-0w7s6qw,jemMS0eyKSdw,Lf	05a8gOVG7UI,AMQrfOzdCms,	190
100007524255117158360	US	GUIRHfx0wC8,VfZhC2zXesc,Z8	(Memo)	9txg0XicoJ0,MsnlBtlmBA,mhHr	134
100007787736839028580	US	I_P_dF9BNkc,	_7hkG6OSo3E,lPnY2NjSjrg,U6,		67
100008764570248830455	US	74tSs7Q3UD8,aFA5q1GNn4w,	_wxw-Z3jic,07i8WF1tAFM,0n)		19
10001231718441705017	VI		(Memo)	BVU7VBO7AZc,mRvkikVuojU,dk	279
100013687611442258419	GB	dg-OCNu_9qc,i4H2UMSCvB0,r		TtBikHrrrls,x3whk3eOuLQ,dg-c	24
100016383867666174158	FI	ERv2SOGL2yw,YWdyeloi_tU,i	vPUqQrMEs2Y,ViZBwbzOcC8,X	o6G7oPA37dM,Xr0V_fiPBuY,pv	207
100018122474852895346	CA	ZXH2M4GbpY0,1UjnBLAnAM,e	QrJ_X82bRoA,pUZQYUahMWC		29
100018280541721034027	NL		0xaPuxZbmUk,A0arz3vb6sA,D	oAnLNx_SFr4,sB1mE8e35UY,Z	45
100019916313199822340	US	Wv4WbzeUeB8,divwIYHdEXI,y	187GeEzMhD4,lf5E8wxZQLw,el		37
100021071340321200633	US		KxWMbqxgMDo,eIRe9uKLjUk,	giRLUIkf4Y,XbZnCGyPneo,GP	107
100022733497715435735	US	RSQ6M4vPdT,A,veIRVnUepYk,	s6NDY8FSr9M,MOcy4sjwcuY,lt	4fa7t_glzgc,Po7-kuw7ziY,gAqf	165
10002320299560720455	US	fzFGchj5K6I,WvuUthifeVs,Afw	9QHc8nZuyaM,yEc2kDXH9DA,		23
100024084709537014786	US	FZadCgdTDmA,ucNeyQi55Zc,j	flkgNn50k14,kfx6zb6Upc,ggV	epW2j86VDwA,jduy12z26s,hf	91
100024316682214673602	CA		rV6SmY04WdE,-5uzJvkealI,y!	1D1cap6yETA,evsttffaWvY,ev	165
100030473591634401277	IT	sXYVWIQyMD8,F1Bm-CEpnGw	bBx2Y5HhpI,arMy72ntu7I,JuN		27
100030967570283228843	US		(Memo)	rT_OmTMwvZI,Mn_StnQVGSo,	35
100033888377343212705	FR	_fk8XtoLjtA,pQou3rgYEAs,XI	pQou3rgYEAs,bLNxHW4ZgXE,	JwQZQygg3Lk,Euci0_BBmNE,rr	97
100034066661652071873	US	Nrk01v9sQC4,9dN-DNvYIAk,5	6W07bFa4TzM,sIFYPQjYhv8,E	zfIkdbWwruY,yxschLOAr-s,ddt	29
100034081581427776311	DK	pJA1CWZl87g,q7d7plCkxJ0,P1	(Memo)	QnjRrTPr1lQ,C_zwJcQ1HO0,Z	24
100034580744319512618	CA	HoXCXc9K4rY,C-T5jk8f1E,uVA	kPRA0W1kEcG,NRdxOr-uYIu,I	CcXXQ6GClb8,vUj26Sybu8c,A	157
100034828195835575273	GB	NGOw9J8zYNA,LjktGpen7w,G	(Memo)	GrFO24ta8J4,WvRWEagxJ5M,	180
100035762233109552669	US	boh0ThiYyhU,_CCWv6EyIeI,J	ffv-lLe0RCw,TjOobAmOe0E,Hr	_CCWv6EyIeI,JXhDEQx0Xmk,	339

Figure 2: A snapshot of YouTube user info table

Each row represents one user, ‘uploadedVideos’, ‘favoriteVideos’ and ‘playListVideos’ contain all the related video ids for the user’s uploads, favorite videos and playlist videos, respectively. Note that the video ids are separated by the ‘,’ separator.

Table 4 – YouTube video metadata table (*youtubevideoinfo.sql*)

This table contains all the textual metadata of the related videos for the users in table 3, including video title, category, tags, description, uploaded time, etc. A snapshot of this table is presented as below:

videoid	title	category	tags	description	dateUploaded	duration	rating	views
--a4AU0MqbI	No Problem	music	no, problem, my, idea, not, big, My Idea	2008-08-06 20:39	130	4.91111	125207	
--Bl_Kr1VU	Alt 0160	film	stop, motion, fotofrafia	2007-08-06 21:58	159	5	1585	
--jB79wLJk	Water heater thermo coupling	howto	ace hardware, true value, plar	2010-12-06 02:20	421	4.75758	27662	
--k2PvSSCE	Hakan does a backflip in Lisbon	comedy	stunt	Hakan decides to be crazy wht	2011-03-15 19:43	17	5	63
--rocYf0-8	Dr. Khans 2007 Entrance	entertainment	XWA, wrestling, Xbox, LIVE, 3	This is Doc Khans XWA 2007 Er	2006-12-07 19:30	76	5	112
--yUDGezQ	Rainy Day Gteborg, 1912	travel	8mm, iPhone, nexvio	Shot with 8mm App on my iPho	2012-05-12 00:55	11	0	14
--0AD0yvuo	HIGH ON FIRE - "10000 Years"	music	high on fire, holderbeast, mikel	HIGH ON FIRE performing "10C	2012-11-24 05:16	307	4.90909	9756
--0CCaRzK4	"Branino kolo" - Sandra Milosevic	music	Sandra Milosevic, Branimir Djok	U emisija "Zikina sarenica", Sanj	2011-11-07 21:46	193	4.88095	110972
--0AvgpjQU	Taito F3 - Puzzle Bobble 2 (Bad	gaming	arcade game, games, video ga	I got this Taito F3 for free, but	2012-11-09 11:41	490	5	412
--0sentorgY	Arena Warrior & Frost DK 2x2	gaming	arm, warrior, pvp, arena, wow	Dévils, and Zrcriticals doing soi	2011-07-26 06:03	349	5	294
--0vedoudq8	Alice Cooper - Identity Crises	music	alice, cooper, identity, crises	Alice Coopers music video for "	2010-09-23 04:01	170	5	6930
--0yETrgf7Y	River of Time 2007 - Bay City,	education	bay city, michigan, rver of tim	2007-The River of Time is spo	2010-07-06 16:18	176	0	147
--0_velNR7sI	Sixty seconds of serenity	entertainment	mobile	This video was uploaded from	2011-06-19 04:26	61	0	9
--1dHDyxG	360°: Ottoneum in Kassel - die	science	Museum, Carl, Schildbach, Hol	Im Naturkundemuseum in Kass	2009-10-03 18:23	31	0	486
--1EyyxtA5w	Wulfgar Invasion? 28/03/2010 -	gaming	wulfgar, invasion, portland, et	Wulfgar na invas?o de Portlan	2010-03-28 23:27	544	5	409
--1oxBW0Mek	Rabbi Ovadia Yosef's Funeral	news	P news	funeral, Jerusalem CityTownVII Rabbi Ovadia Yosef, 93, one o	2013-10-07 22:31	240	5	983
--1pa7gw15g	Opium - AVROs Jazzarchief - R			Ray Charles, jazzarchief, AVR	De AVRO heeft een rijk archief	2012-02-17 21:30	380	5
--1TaQerft	Buenafuente (915) - Pau Gas	entertainment	buenafuente, informatica, eye	Entrevista de Buenafuente a P	2011-04-08 17:01	1371	4.90184	34806
--1Uj8E-ipA	Actor Richard Kind Talks TV & I	news	Richard Kind, actor, television,	From the Majority Report, live	2012-04-13 17:29	1319	4.78947	562
--1W3Kq_nPr	Triston Palmer - Joker Smoker	music	triston, palmer, tristan, palma	Label: Midnight Rock	2008-02-06 13:17	216	4.95202	357414
--1y1Qb0i64	Carbonite - Gather Routing	gaming	WoW Marcos Guide, WoW leve	WoW Marcos Guide power leve	2011-05-28 09:01	441	3.66667	425
--22yFZGIGA	Dions 7th birthday	people	pixelpipe	via Pixelpipe	2010-11-06 19:22	32	0	8
--28UKighL8	Cold freaks No homo	people	Freaks	something for the ladies	2008-01-11 21:14	237	3.68085	171945
--2a3AEPk9U	World Champion Cyclo Cross 2	sports	world, champion, cyclo, cross,	For one moment I thought my	2007-01-29 21:41	60	4.79487	71549
--2c9pegg0M	Wont Get Fooled Again - Anim	entertainment	The, Who, Wont, Get, Fooled,	An animation I made using bon	2007-06-23 00:21	256	4.45098	4505
--2CNlelQ	Duck Sauce - Big Bad Wolf [HD	music	Duck Sauce, Big Bad Wolf, Unc	Join us on Spinnin Facebook: ?	2011-09-29 11:48	181	4.3703	258489
--2dylyYPHE	Productionalize,maybe?	education	vidcat2	efforts in streamlining our proc	2012-07-31 14:08	566	0	63
--2GxKmLJ	Shazida Khatun has Captured	entertainment	kalgachia, shazida khatun, goe	shazid Khatun has captured th	2010-07-17 10:20	283	5	671
--2T0GQ-Q	Love Games between Dog and pets		Hund, Katze, dog, cat, pet, gei	der ?berzeugungseffekt ist so	2012-03-31 21:52	43	5	64
--2NdcywRju	Alex Lifeson's Guitar Tech - Scr	music	Rush, Alex Lifeson, Harmony	C Scott Appleton, Guitar Tech fo	2011-04-29 22:33	596	4.94406	47375
--2t1r52pp4	???? ??????? ?????? ??????	music	, 2	???? ??????? ????????	2011-12-08 17:23	60	0	178
--3190nZzk	Doutzen Kroes Part 1/3 English	entertainment	Doutzen, Kroes, Documentary	/ Documentary about topmodel I	2008-01-14 19:31	561	4.7861	130594
--3eIR5g2R1	Paul van Dyk - For An Angel (1	entertainment	electronic, Paul Van Dyk, Disc	J Paul van Dyk - For An Angel (1	2009-10-01 18:28	232	4.97552	260411
--3gt3sys-M	F.E.A.R. - Few scares	gaming	FEAR, pc gaming	A few semi scary moments in F	2011-04-03 23:13	56	0	17
--3ISpA5FY	Skier Cheats Avalanche With E	entertainment	Skier, Cheats, Avalanche, Epic	Click here to subscribe now:	2013-02-14 12:00	119	5	467
--3LuiMIVSQ	Smashing Pumpkins - G.L.O.W.	music	glow, smashing, pumpkins, mu	Poop	2008-12-02 02:55	203	4.85714	261258

Figure 3: A snapshot of YouTube video metadata table

'videoid' uniquely identify the typical video, which can be found in the video list of the table 1, and different tags of each video are separated by the ',' separator in 'tags' attribute.